Mythos Readiness Checklist

Are you prepared to defend against fast-moving AI-powered adversaries?

In April 2026, Anthropic disclosed Claude Mythos, showing how quickly AI is changing the speed and scale of cyberattacks. Advanced models can help attackers identify and exploit weaknesses far faster than traditional security processes were built to handle.

That raises the stakes for common security defensive gaps like weak MFA, excessive permissions, stale firewall rules, missed patches, and configuration drift. These issues are not new, but AI-powered attackers can find and exploit them much faster and at far greater scale.

Security teams now need to close these gaps faster than attackers can exploit them. This checklist helps assess whether your controls are properly configured, enforced, and continuously validated to keep AI-powered adversaries at bay.

DIFFERENT ATTACK PATHS. ONE RECURRING PATTERN.

PHISHABLE MFA AND HELP-DESK RESETS

Groups like Scattered Spider compromise identity by calling help desks, impersonating employees, and talking their way into a password or MFA reset. AI-assisted pretexting makes these calls faster to script and harder to catch.

The opening: phishable MFA, weak reset verification, and no monitoring of new device enrollments.

OVER-SCOPED OAUTH AND SAAS INTEGRATIONS

ShinyHunters-linked campaigns rarely use malware. They reuse valid OAuth tokens and connected-app permissions to pull data straight out of Salesforce, Workday, and similar platforms at scale.

The opening: broad, unreviewed API scopes and connected apps with standing access nobody revisits.

UNPATCHED NETWORK EDGE DEVICES

Firewalls, VPN concentrators, and remote-access gateways are attractive because a single misconfigured or unpatched device can expose an entire network, and patch cycles for edge hardware routinely lag behind the rest of the stack.

The opening: stale firewall rules, internet-facing management interfaces, and delayed edge patching.

DEEPFAKE-ENABLED SOCIAL ENGINEERING

AI-generated deepfakes in business email compromise grew from under 5% of cases in 2023 to roughly 40% by early 2026. In one widely documented case, a finance employee authorized around $25 million in fraudulent transfers after a video call where every other participant was an AI-generated deepfake.

The opening: no callback or out-of-band confirmation step for high-risk approvals and payment changes.

Score your security controls

Check each box only if the statement is true and enforced in your production environment today.

IDENTITY AND MFA HARDENING

- [ ] Administrative and sensitive accounts require phishing-resistant MFA, such as FIDO2 or passkeys, not SMS or voice codes.
- [ ] Enrolling a new MFA factor requires strong identity verification, not a phone call alone.
- [ ] Help-desk password and MFA resets follow a documented process that resists impersonation.
- [ ] Dormant, contractor, and unnecessary local accounts are identified and removed on a regular schedule.

SAAS AND OAUTH ACCESS

- [ ] Connected applications and OAuth integrations are reviewed on a set schedule, not left unchecked indefinitely.
- [ ] API tokens and sessions expire or rotate on a defined interval, with stale grants revoked automatically.
- [ ] Bulk exports and unusual download volume can be detected, investigated, and blocked in real time.
- [ ] Third-party app permissions are scoped to least privilege rather than broad, standing access.

NETWORK AND EDGE EXPOSURE

- [ ] Internet-facing administrative and management interfaces are continuously inventoried.
- [ ] Firewall, VPN, and access policies are reviewed for stale rules or unnecessarily broad access.
- [ ] Patches for edge devices are applied against a defined SLA rather than on an ad hoc basis.
- [ ] Changes that broaden network access are reviewed before they reach production.

AI-ERA VALIDATION AND DRIFT

- [ ] Conditional access policies flag or block sign-ins from new devices, unfamiliar locations, or impossible travel for finance and executive accounts.
- [ ] Changes to MFA, permissions, guest access, APIs, and network policy generate alerts as they happen.
- [ ] Actual configurations are compared against an approved security baseline continuously, not only at audit time.
- [ ] Control gaps have a named owner and a remediation deadline measured in hours, not weeks.

YOUR SCORE

0–6 CHECKED

You are significantly exposed. AI-speed attackers find and chain these gaps in minutes, not days.

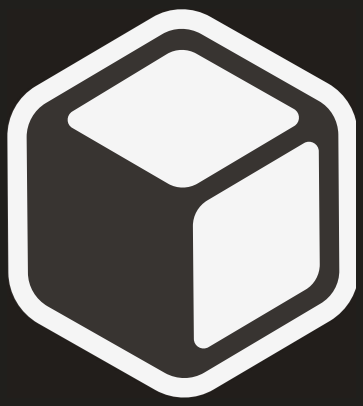
7–13 CHECKED

Some protection is in place, but AI-accelerated reconnaissance will still find the paths left open.

14–16 CHECKED

Your controls are well configured. Validate them continuously as environments and threats change.

Close security control gaps before AI attackers find them



Security controls change constantly. Permissions expand, exceptions stay in place longer than intended, capabilities get disabled during patch cycles, and policies drift away from the design security teams intended.

Reach gives security teams continuous visibility into how controls are actually configured and whether those controls remain aligned, enforced, and ready to stop fast-moving AI-powered adversaries.

By analyzing the tools already deployed across your environment, Reach finds and fixes misconfigurations across your security and IT stack, activates underused protections, detects drift and corrects it, and closes defensive gaps faster than AI-powered attackers can find and exploit them.

01

FIND THE GAPS THAT CREATE EXPOSURE

Identify missing protections, overly permissive policies, ineffective enforcement, phishable MFA, and configurations that have drifted from their approved state.

02

PRIORITIZE WHAT TO FIX FIRST

Focus remediation on the gaps that create the most exposure given your environment, active threats, and the controls already available to you.

03

AUTOMATICALLY REMEDIATE

Give teams clear, step-by-step and automated guidance to correct each issue and bring controls back into alignment with policy at speed and scale.

04

CONTINUOUSLY VALIDATE

Monitor configurations to detect drift the moment it happens, correct it, and confirm controls are working as intended to keep AI-powered attackers at bay.

Security controls only protect you when they're configured correctly and validated continuously.

Reach helps you find and fix control gaps faster than AI-fueled attackers can find and exploit them.

Request a Demo

reach.security

Research sources: Anthropic, Project Glasswing and Claude Mythos Preview announcements (April 2026); UK AI Security Institute, evaluation of Claude Mythos Preview's cyber capabilities; CyberScoop and Cybersecurity Dive, Project Glasswing expansion coverage (June 2026); Computer Weekly, "Cyber protection against advances in frontier AI models"; The Guardian, "What is Mythos AI and why could it be a threat to global cybersecurity?" (April 22, 2026); CrowdStrike 2026 Global Threat Report; Verizon 2025 Data Breach Investigations Report; FBI Internet Crime Complaint Center, 2025 Internet Crime Report; Reach Security internal research on security control misconfiguration.